CLAIMS

What is claimed is:

1. A method for evaluating a plurality of candidate index sets for a workload of database statements in a database system, the method comprising:

generating baseline statistics for each statement in the workload;

forming an index superset from a union of a current index set and a proposed index set;

deriving a candidate index set from the index superset, the candidate index being one of the plurality of candidate index sets;

generating statistics based on the candidate index set and the baseline statistics; and

presenting the generated statistics.

2. The method of Claim 1, further comprising:

generating current index statistics for the workload responsive to the current index set, the presented generated statistics comprising the generated current index statistics.

3. The method of Claim 1, further comprising:

repeatedly deriving a candidate index set and generating statistics based on the proposed index set.

4. The method of Claim 3, further comprising:

terminating the repeated execution when at least one candidate index solution is found that adheres to user-imposed constraints and no further indexes can be removed from said candidate index solution without degrading performance of the workload and without disabling an integrity constraint.

5. The method of Claim 1, wherein deriving the baseline statistics comprises disabling current indexes.

6. The method of Claim 1, wherein generating statistics for a statement comprises:

creating an execution plan which represents a series of steps for executing the statement;

evaluating the execution plan;

generating and recording statistics based on the evaluation of the execution plan.

7. The method of Claim 6, wherein creating an execution plan is based on available access paths.

8. The method of Claim 6, wherein creating an execution plan is based on statistics for at least one schema object accessed by the statement.

9. The method of Claim 8 wherein the at least one schema object is a table.

10. The method of Claim 8 wherein the at least one schema object is an index.

11. The method of Claim 6, wherein evaluating the execution plan comprises:

for a table accessed by a statement under evaluation, identifying at least one index that would be used to retrieve data from the table upon an execution of the statement.

12. The method of Claim 6, wherein evaluating the execution plan comprises:

determining a cost of the execution plan.

13. The method of Claim 12, wherein the cost of the execution plan is derived from a resource use needed to execute the statement according to the execution plan.

14. The method of Claim 13, wherein the resource use includes CPU execution time.

15. The method of Claim 13, wherein the resource use includes input/output access.

16. The method of Claim 6, wherein the statistics include the number of executions of the statement.

17. The method of Claim 6, wherein the statistics include a user-defined importance of the statement.

18. The method of Claim 6, wherein the statistics include an index usage.

19. The method of Claim 6, wherein the statistics include a cost of the execution plan.

20. The method of Claim 1, wherein the statements are SQL statements.

21. The method of Claim 1, wherein the workload is reduced into unique statements.

22. The method of Claim 1, wherein deriving a candidate index set is responsive to a predetermined maximum number of allowed indexes.

23. The method of Claim 1, wherein deriving a candidate index set is responsive to available storage space.

24. The method of Claim 1, wherein the proposed index set is provided by a user.

25. The method of Claim 1, wherein the proposed index set is provided by an expert system.

26. The method of Claim 1, wherein an execution plan is created without creating indexes which are not in the current index set.

27. A system for evaluating a plurality of candidate index sets for a workload in a database system, the workload derived from a plurality of statements, the system comprising:

a workload evaluator which evaluates each statement within the workload;

an index solution evaluator which, responsive to the workload evaluator, evaluates each index in a candidate index set with respect to the workload, the candidate index solution being one of the plurality of candidate index sets;

a solution/rollup evaluator which, responsive to the index solution evaluator, evaluates the candidate index solution; and

a solution refiner which, responsive to the solution/rollup evaluator, generates at least one new candidate index solution.

28. The system of Claim 27, wherein the solution refiner generates at least one new candidate index solution by eliminating at least one index within the candidate index solution that does not adhere to user-imposed constraints.

29. The system of Claim 28, wherein the constraint is a user-defined constraint.

30. The system of Claim 28, wherein the constraint is a memory-usage constraint.

31. The system of Claim 27, wherein the solution refiner generates at least one new candidate index solution by eliminating at least one index on a small table under evaluation, and wherein the at least one index does not enforce an integrity constraint.

32. The system of Claim 27, wherein the evaluation created by the workload evaluator comprises an execution plan for each statement which represents a series of steps for executing the statement, wherein the workload evaluator evaluates the execution plan, and generates and records statistics based on the evaluation of the execution plan.

33. The system of Claim 32, wherein each execution plan is created based on available access paths.

34. The system of Claim 32, wherein each execution plan is created based on statistics for at least one schema object accessed by the statement.

35. The system of Claim 34 wherein the at least one schema object is a table.

36. The system of Claim 34 wherein the at least one schema object is an index.

37. The system of Claim 32, wherein the workload evaluator, for a table accessed by a statement under evaluation, identifies at least one index which would be used to retrieve data from the table upon an execution of the statement.

38. The system of Claim 32, wherein the workload evaluator determines a cost of the execution plan.

39. The system of Claim 38, wherein the cost of the execution plan is derived from a resource use needed to execute the statement according to the execution plan.

40. The system of Claim 39, wherein the resource use includes CPU execution time.

41. The system of Claim 39, wherein the resource use includes input/output access.

42. The system of Claim 32, wherein the statistics include the number of executions of the statement.

43. The system of Claim 32, wherein the statistics include a user-defined importance of the statement.

44. The system of Claim 32, wherein the statistics include an index usage.

45. The system of Claim 32, wherein the statistics include a cost of the execution plan.

46. The system of Claim 27, wherein the statements are SQL statements.

47. The system of Claim 27, wherein the workload is reduced into unique statements.

48. The system of Claim 27, wherein the solution refiner is responsive to a predetermined maximum number of allowed indexes.

49. The system of Claim 27, wherein the solution refiner is responsive to available storage space.

50. The system of Claim 27, wherein the proposed index set is provided by a user.

51. The system of Claim 27, wherein the proposed index set is provided by an expert system.

52. The system of Claim 27, wherein an execution plan is created without creating indexes which are not in the current index set.

53. A computer program product for evaluating a plurality of candidate index sets for a workload of database statements in a database system, the computer program product comprising a computer usable medium having computer readable code thereon, including program code which:

generates baseline statistics for each statement in the workload;

forms an index superset from a union of a current index set and a proposed index set;

repeatedly

derives a candidate index set from the index superset, the candidate index superset being one of the plurality of candidate index sets, and

generates statistics based on the candidate index set and the baseline statistics; and

presents the generated statistics.

54. A computer data signal embodied in a carrier wave for evaluating a plurality of candidate index sets for a workload of database statements in a database system, comprising:

program code for generating baseline statistics for each statement in the workload;

a6 cont

program code for forming an index superset from a union of a current index set and a proposed index set;

program code for repeatedly

deriving a candidate index set from the index superset, the candidate index superset being one of the plurality of candidate index sets, and

generating statistics based on the candidate index set and the baseline statistics; and

program code for presenting the generated statistics.